MODELS ALLOWING FOR AGE-DEPENDENT SURVIVAL RATES FOR BAND-RETURN DATA

by C. Brownie and D. S. Robson
Biometrics Unit, Cornell University, Ithaca, N.Y.

BU-514-M June, 1974

Summary

Seber (1970) and Robson and Youngs (1971) developed a statistical model for a tag-recapture experiment on an exploited animal population in which banded or tagged individuals are released into a population each year, and the bands from harvested individuals are returned by hunters or fishermen in following years. Consideration of certain biological and ecological factors affecting bird populations has led to the proposal of new models for an experimental situation where data are recorded separately for birds released as adults and as young-of-the-year. In particular, the effect of different survival rates of young birds, and the effect of migration on reporting rates, are taken into account in the models presented in this study. Maximum likelihood estimators of parameters are derived under the different models, and procedures are developed to discriminate between alternative models as well as provide non-discriminant goodness of fit tests.

1. Introduction

In the type of banding experiment considered here, a known number of banded birds is released into the population each summer for a number of years. The population is harvested during the hunting season each year, bands from captured birds are returned by hunters, and a yearly record is kept of the number of returns from each batch released. Seber (1970) and Robson and Youngs (1971) developed a stochastic model for this situation under the assumption that annual survival, exploitation and reporting rates are year specific but independent of age. This model, (which we shall refer to as the SRY model), is a special case of the more general tag-recapture model given by Jolly (1965), and the SRY estimators can be derived from the Jolly estimators. For reasons outlined below, the assumptions made under the SRY model, with respect to annual survival, exploitation and reporting rates are probably not realistic. This has lead to the development of new models, which are sufficiently different from the Jolly model to merit separate consideration.

The assumption that survival rates are independent of age, restricts application of the SRY model to data from birds banded as adults only since young and adult birds are thought to have different survival rates. However if records of releases and returns are kept separately for birds banded as adults and as young-of-the-year, then models can be employed which admit different survival and reported exploitation rates of young birds.

Another limitation of the SRY model is the assumption that reporting rates are constant for all birds in a given year. Reporting rates have been noted to be lower near the banding site, probably because hunters there are more accustomed to seeing bands and so return them at a lower rate. After migration birds do not necessarily return to the banding site, and hence the reporting rate for birds banded and released in a given year should be different from that for birds banded in previous years since the latter have undergone at least one migration. This leads to another modification of the SRY model.

The development of models reflecting the above considerations, corresponding estimation formulae, and certain tests of hypothesis follow.

2. The Model Under H_1

In general it is possible to distinguish only two age-groups of birds, namely young-of-the-year birds and adults (i.e., birds more than one year old). Thus consider a banding experiment where known numbers of adults and young birds are released into the population each year before the hunting season, and yearly records of band returns are kept separately for birds banded and released as adults and for birds banded and released as young-of-the-year.

Consider the i^{th} year of the banding experiment to be the period between the i^{th} and $(i+1)^{th}$ banding dates. Let N_i and M_i be the number of adult and young birds respectively which are banded and released into the population in year i, $i=1,\ldots,k$.

Let R_{ij} be the number of bands returned in year j from the batch of adults released in year i, $i=1,\ldots,k$, $j=i,\ldots,k$.

Let Q_{ij} be the number of bands returned in year j from the batch of young birds released in year i, $i=1,\ldots,k$, $j=i,\ldots,k$.

Let $\{R_{ij}\}$ be the triangular array of random variables R_{ij}, i.e.,

$$\{R_{ij}\} = \begin{Bmatrix} R_{11} & R_{12} & \cdots & R_{1k} \\ & R_{22} & \cdots & R_{2k} \\ & & \ddots & \vdots \\ & & & R_{kk} \end{Bmatrix}$$

Similarly let $\{Q_{ij}\}$ be the triangular array of random variables Q_{ij}.

For any banded adult alive at the start of year i, we define the following conditional probabilities:

S_i = P [survives year i] = survival rate.

u_i = P [killed by hunter in year i] = exploitation rate.

λ_i = P [band returned given banded adult killed by hunter in year i] = reporting rate.

$f_i = \lambda_i u_i$ = P [killed by hunter and band returned in year i] = reported exploitation rate.

Similarly, for any young bird banded and released at the start of year i, let

S_i' = P [survives year i].

f_i' = P [killed by hunter and band returned in year i].

Thus, as in the SRY model, survival, exploitation and reporting rates are assumed to be year specific, but in addition it is assumed that survival and exploitation rates are different for young-of-the-year and adult birds, and we define H_1 to be the hypothesis that this last assumption is true. H_1 can be characterized by the following.

$$E_{H_1}[\{R_{ij}\}] = \begin{Bmatrix} N_1f_1 & N_1S_1f_2 & \cdots & N_1S_1\cdots S_{k-1}f_k \\ & N_2f_2 & \cdots & N_2S_2\cdots S_{k-1}f_k \\ & & \ddots & \vdots \\ & & & N_kf_k \end{Bmatrix}$$

$$E_{H_1}[\{Q_{ij}\}] = \begin{Bmatrix} M_1f_1' & M_1S_1'f_2 & \cdots & M_1S_1'S_2\cdots S_{k-1}f_k \\ & M_2f_2' & \cdots & M_2S_2'S_3\cdots S_{k-1}f_k \\ & & \ddots & \vdots \\ & & & M_kf_k' \end{Bmatrix}$$

Assuming that banded birds released into the population suffer statistically independent fates, the likelihood under H_1 of $\{R_{ij}\}$ is the same as that derived by Robson and Youngs (1971), i.e.,

$$P_{H_1}[\{R_{ij}\}] = \prod_{i=1}^{k} \binom{N_i}{R_{ii},\ldots,R_{ik}} f_i^{R_{\cdot i}} \, S_{i-1}^{T_i - R_{i\cdot}} \, (1-\rho_i)^{N_i - R_{i\cdot}}$$

where $R_{i\cdot}, R_{\cdot i}$ are row and column totals respectively of the array $\{R_{ij}\}$, $T_1 = R_{1\cdot}$, $T_{i+1} = R_{i+1\cdot} + T_i - R_{\cdot i}$, $i=1,\dots,k-1$, and

$$\rho_i = E_{H_1}\left[\frac{R_{i\cdot}}{N_i}\right] = \begin{cases} f_i + S_i f_{i+1} + \dots + S_i S_{i+1} \dots S_{k-1} f_k & i=1,\dots,k-1 \\ f_k & i=k \end{cases}$$

and f_i, S_i, ρ_i correspond respectively to Seber's $\alpha_i, \beta_i, \theta_i$.

Similarly defining $Q_{i\cdot}, Q_{\cdot i}$ to be the row and column totals respectively of the array $\{Q_{ij}\}$, and $U_1 = Q_{1\cdot}$, $U_{i+1} = Q_{i+1\cdot} + U_i - Q_{\cdot i}$, $i=1,\dots,k-1$, and

$$\rho_i' = E_{H_1}\left[\frac{Q_{i\cdot}}{M_i}\right] = \begin{cases} f_i' + S_i' \rho_{i+1} & i=1,\dots,k-1 \\ f_k' & i=k \end{cases}$$

we have

$$P_{H_1}[\{Q_{ij}\}] = \prod_{i=1}^{k} \binom{M_i}{Q_{ii},\dots,Q_{ik}} f_i'^{\,Q_{ii}} \, S_i'^{\,Q_{i\cdot}-Q_{ii}}$$

$$f_i^{\,Q_{\cdot i}-Q_{ii}} \, S_{i-1}^{\,U_i - Q_{i\cdot} - Q_{i-1\cdot} + Q_{i-1,i-1}}$$

$$(1-\rho_i')^{M_i - Q_{i\cdot}}$$

where $Q_{0,i} \equiv 0$. Hence

$$P_{H_1}[\{R_{ij}\},\{Q_{ij}\}] = \prod_{i=1}^{k} \binom{N_i}{R_{ii},\dots,R_{ik}} \binom{M_i}{Q_{ii},\dots,Q_{ik}} f_i^{W_i} \, S_i^{Z_{i+1}}$$

$$f_i'^{\,Q_{ii}} \, S_i'^{\,Q_{i\cdot}-Q_{ii}} \, (1-\rho_i)^{N_i - R_{i\cdot}} \, (1-\rho_i')^{M_i - Q_{i\cdot}}$$

where $\quad W_i = R_{\cdot i} + Q_{\cdot i} - Q_{ii} \qquad i=1,\dots,k$

$$Z_i = \begin{cases} T_i + U_i - R_{i\cdot} - Q_{i\cdot} - Q_{i-1\cdot} + Q_{i-1,i-1} & i=2,\dots,k \\ 0 & i=k+1 \end{cases}$$

By expressing the likelihood in this form, and noting that $T_{i+1}+U_{i+1} = R_{i+1.}+Q_{i+1.}+T_i+U_i-W_i-Q_{ii}$, $i=2,\ldots,k-1$, we can identify a minimal sufficient statistic $\mathscr{S}_1$ under H_1 as

$$\mathscr{S}_1 = \{R_{1.},\ldots,R_{k.},Q_{1.},\ldots,Q_{k.},Q_{11},\ldots,Q_{k-1,k-1},T_2+U_2,\ldots,T_k+U_k\}$$

Now note that the likelihood can be expressed as the product

$$P_{H_1}[\{R_{ij}\},\{Q_{ij}\}] = P_{H_1}[\mathscr{S}_1]\; P_{H_1}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_1]$$

$$= \prod_{i=1}^{k-1} \{P_{H_1}[R_{i.}|N_i]\; P_{H_1}[Q_{i.}|M_i]\; P_{H_1}[Q_{ii}|Q_{i.}]$$

$$P_{H_1}[W_i|W_i+Z_{i+1}]\}\quad P_{H_1}[R_{k.}|N_k]$$

$$P_{H_1}[Q_{k.}|M_k]\; P_{H_1}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_1]\ ,$$

where the residual distribution given by $P_{H_1}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_1]$ is independent of the unknown parameters, and

$$R_{i.} \sim \text{Bin}\ (N_i,\rho_i), \qquad i=1,\ldots,k$$

$$Q_{i.} \sim \text{Bin}\ (M_i,\rho_i'), \qquad i=1,\ldots,k$$

$$Q_{ii} \text{ given } Q_{i.} \sim \text{Bin}\ (Q_{i.},f_i'/\rho_i'), \qquad i=1,\ldots,k-1$$

$$W_i \text{ given } W_i+Z_{i+1} \sim \text{Bin}\ (W_i+Z_{i+1},f_i/\rho_i), \qquad i=1,\ldots,k-1$$

and Bin (N,ρ) represents the Binomial distribution with parameters N,ρ. Therefore, the maximum likelihood estimators of the parameters ρ_i,ρ_i', f_i/ρ_i, f_i'/ρ_i' are:

$$\hat{\rho}_i = \frac{R_{i.}}{N_i}\ , \quad \hat{\rho}_i' = \frac{Q_{i.}}{M_i}\ , \qquad i=1,\ldots,k,$$

$$\frac{\hat{f}_i}{\rho_i} = \frac{W_i}{W_i+Z_{i+1}}\ , \quad \frac{\hat{f}_i'}{\rho_i'} = \frac{Q_{ii}}{Q_{i.}}\ , \qquad i=1,\ldots,k-1,$$

and using the relationships $\rho_i = f_i + S_i\rho_{i+1}$, $\rho'_i = f'_i + S'_i\rho_{i+1}$, gives the maximum likelihood estimators of the parameters f_i, S_i, f'_i, S'_i:

$$\hat{f}_i = \frac{R_{i\cdot}}{N_i}\,\frac{W_i}{W_i+Z_{i+1}}, \qquad \hat{f}'_i = \frac{Q_{ii}}{M_i}, \qquad i=1,\ldots,k$$

$$\hat{S}_i = \frac{\dfrac{R_{i\cdot}}{N_i} - \hat{f}_i}{\dfrac{R_{i+1\cdot}}{N_{i+1}}}, \qquad \hat{S}'_i = \frac{Q_{i\cdot}-Q_{ii}}{M_i}\,\frac{N_{i+1}}{R_{i+1\cdot}}, \qquad i=1,\ldots,k-1\;.$$

The estimators $\hat{f}_i$ can be shown to be unbiased, and the estimators $\hat{S}_i, \hat{S}'_i$ are consistent.

<u>Variances and covariances of maximum likelihood estimators under H_1</u>

$$\mathrm{Var}\,(\hat{f}'_i) = f_i(1-f_i)/M_i, \qquad i=1,\ldots,k$$

Using conventional Taylor series methods, the following asymptotic variances and covariances are obtained:

$$\mathrm{Var}\,(\hat{f}_i) \doteq f_i^2\left[\frac{1}{E(R_{i\cdot})} - \frac{1}{N_i} + \frac{1}{E(W_i)} - \frac{1}{E(W_i+Z_{i+1})}\right], \qquad i=1,\ldots,k,$$

$$\mathrm{Var}\,(\hat{S}_i) \doteq S_i^2\left[\frac{1}{E(R_{i\cdot})} - \frac{1}{N_i} + \frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}} + \frac{1}{E(Z_{i+1})} - \frac{1}{E(W_i+Z_{i+1})}\right],$$

$$i=1,\ldots,k-1,$$

$$\mathrm{Cov}\,(\hat{f}_i,\hat{S}_i) \doteq f_iS_i\left[\frac{1}{E(R_{i\cdot})} - \frac{1}{N_i} - \frac{1}{E(W_i+Z_{i+1})}\right], \qquad i=1,\ldots,k-1,$$

$$\mathrm{Cov}\,(\hat{f}_{i+1},\hat{S}_i) \doteq -S_if_{i+1}\left[\frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right], \qquad i=1,\ldots,k-1,$$

$$\mathrm{Cov}\ (\hat{S}_i,\hat{S}_{i+1}) \doteq -S_iS_{i+1}\left[\frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right], \qquad i=1,\ldots,k-2,$$

$$\mathrm{Cov}\ (\hat{S}'_i,\hat{S}_i) \doteq S'_iS_i\left[\frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right], \qquad i=1,\ldots,k-1,$$

$$\mathrm{Cov}\ (\hat{S}'_i,\hat{S}_{i+1}) \doteq -S'_iS_{i+1}\left[\frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right], \qquad i=1,\ldots,k-2,$$

$$\mathrm{Cov}\ (\hat{S}'_i,\hat{f}_{i+1}) \doteq -S'_if_{i+1}\left[\frac{1}{E(R_{i+1\cdot})} - \frac{1}{N_{i+1}}\right], \qquad i=1,\ldots,k-1,$$

$$\mathrm{Cov}\ (\hat{f}'_i,\hat{S}'_i) \doteq \frac{f'_iS'_i}{M_i} \qquad i=1,\ldots,k-1,$$

where expectations are with respect to the model under H_1.

All other covariances are exactly or approximately zero. Estimators of variances and covariances are obtained in an obvious way by replacing expected values by observed values in the above formulae.

The above derivations apply to the case where the data arrays $\{R_{ij}\},\{Q_{ij}\}$ are triangular, but the extension to the case where the data arrays have k rows and k+s columns, $s > 0$, is straightforward. In this case the minimal sufficient statistic is $\mathscr{S}_1^* = \{R_{1\cdot},\ldots,R_{k\cdot},\ Q_{1\cdot},\ldots,Q_{k\cdot},\ Q_{11},\ldots,Q_{kk},\ T_2+U_2,\ldots,$ $T_k+U_k,\ R_{\cdot k+1}+Q_{\cdot k+1},\ldots,R_{\cdot k+s}+Q_{\cdot k+s}\}$ and the formulae for the estimators $\hat{f}_i,\hat{f}'_i$, $i=1,\ldots,k$, $\hat{S}_i,\hat{S}'_i$, $i=1,\ldots,k-1$, are the same if $R_{i\cdot},R_{\cdot i},Q_{i\cdot},Q_{\cdot i}$ are defined in the obvious way. As indicated by Seber (1970), the parameters S_{k-1+j},f_{k+j} are not separately identifiable for $j=1,\ldots,s$.

3. A Test of H_0 vs H_1

We now define a test of the assumption that young-of-the-year and adult birds have the same survival and reported exploitation rates. If we let H_0 be the hypothesis that this assumption is true, then the model under H_0 is characterized by

$$E_{H_0}\left[\frac{R_{ij}}{N_i}\right] = E_{H_0}\left[\frac{Q_{ij}}{M_i}\right] = \begin{cases} f_i & i=j,\ i=1,\ldots,k \\ S_i \cdots S_{j-1} f_j & j=i+1,\ldots,k \\ & i=1,\ldots,k-1. \end{cases}$$

Under H_0, the model is analogous to the SRY model and the minimal sufficient statistic, derived by Robson and Youngs (1971), is

$\mathcal{S}_0 = \{R_{1.}+Q_{1.},\ldots,R_{k.}+Q_{k.},\ T_2+U_2,\ldots,T_k+U_k\}.$

The conditional distribution under H_0 of $\mathcal{S}_1$ given $\mathcal{S}_0$ can now be used to derive a conditional test $\phi(\mathcal{S}_0,\mathcal{S}_1)$ of H_0 vs H_1 with the property $E_{H_0}[\phi(\mathcal{S}_0,\mathcal{S}_1)|\mathcal{S}_0] = \alpha.$

In order to obtain $P_{H_0}[\mathcal{S}_1|\mathcal{S}_0]$ we need $P_{H_0}[Q_{1.},\ldots,Q_{k.},\ Q_{11},\ldots,Q_{k-1,k-1}|\mathcal{S}_0]$. Note that $P_{H_0}[Q_{i.},Q_{ii}|\mathcal{S}_0] = P_{H_0}[Q_{i.}|R_{i.}+Q_{i.}]\ P_{H_0}[Q_{ii}|Q_{i.},\mathcal{S}_0]$

$$= \frac{\binom{M_i}{Q_{i.}}\binom{N_i}{R_{i.}}}{\binom{M_i+N_i}{R_{i.}+Q_{i.}}}\ \frac{\binom{Q_{i.}}{Q_{ii}}\binom{T_i+U_i-Q_{i.}}{W_i}}{\binom{T_i+U_i}{R_{.i}+Q_{.i}}}$$

since under H_0 $\quad Q_{i.} \sim \text{Bin}\ (M_i,\rho_i)$

$R_{i.} \sim \text{Bin}\ (N_i,\rho_i)$

and Q_{ii} given $Q_{i.} \sim \text{Bin}\ (Q_{i.},f_i/\rho_i)$

W_i given $T_i+U_i-Q_{i.} \sim \text{Bin}\ (T_i+U_i-Q_{i.},f_i/\rho_i)$

$$\therefore\ P_{H_0}[\mathscr{S}_1|\mathscr{S}_0] = \prod_{i=1}^{k} \frac{\binom{M_i}{Q_{i\cdot}}\binom{N_i}{R_{i\cdot}}}{\binom{M_i+N_i}{R_{i\cdot}+Q_{i\cdot}}} \prod_{i=1}^{k-1} \frac{\binom{Q_{i\cdot}}{Q_{ii}}\binom{T_i+U_i-Q_{i\cdot}}{W_i}}{\binom{T_i+U_i}{R_{\cdot i}+Q_{\cdot i}}}$$

The conditional distribution $P_{H_0}[\mathscr{S}_1|\mathscr{S}_0]$ can therefore be represented as a product of hypergeometric distributions, and a chi-square test can be applied in the usual way to the 2 X 2 table corresponding to each hypergeometric term in the product. Thus a single degree of freedom chi-square statistic is obtained from each of the contingency tables

$$\boxed{\begin{array}{ll} Q_{i\cdot} & M_i-Q_{i\cdot} \\ R_{i\cdot} & N_i-R_{i\cdot} \end{array}} \quad i=1,\ldots,k, \quad \text{and} \quad \boxed{\begin{array}{ll} Q_{ii} & Q_{i\cdot}-Q_{ii} \\ W_i & T_i+U_i-Q_{i\cdot}-W_i \end{array}} \quad i=1,\ldots,k-1.$$

These chi-square statistics are asymptotically independent and may be added to give a single test statistic which is asymptotically chi-square under H_0 with 2k-1 degrees of freedom.

When the data arrays have k rows and k+s columns, s > 0, the sufficient statistic under H_0 is $\mathscr{S}_0^* = \{R_{1\cdot}+Q_{1\cdot},\ldots, R_{k\cdot}+Q_{k\cdot}, T_2+U_2,\ldots,T_k+U_k, R_{\cdot k+1}+Q_{\cdot k+1},\ldots,R_{\cdot k+s}+Q_{\cdot k+s}\}$, and the corresponding test is based on the distribution given by

$$P_{H_0}[\mathscr{S}_1^*|\mathscr{S}_0^*] = \prod_{i=1}^{k} \frac{\binom{M_i}{Q_{i\cdot}}\binom{N_i}{R_{i\cdot}}}{\binom{M_i+N_i}{R_{i\cdot}+Q_{i\cdot}}} \frac{\binom{Q_{i\cdot}}{Q_{ii}}\binom{T_i+U_i-Q_{i\cdot}}{W_i}}{\binom{T_i+U_i}{R_{\cdot i}+Q_{\cdot i}}}$$

4. The Model Under H_2

For reasons discussed in the introduction, another model which is of interest is one under which the assumptions about survival and exploitation rates are the same as under H_1, but in addition it is assumed that in any given year, the reporting rate for birds banded and released in that year is different from that for birds banded and released in previous years. We let H_2 be the hypothesis that these assumptions hold.

Let $f_i''' = P$ [adult released in year i is killed by hunter and its band returned in year i], $i=1,\ldots,k$

and

$$\rho_i''' = E_{H_2}\left[\frac{R_{i\cdot}}{N_i}\right] = \begin{cases} f_i'''+S_i\rho_{i+1} & i=1,\ldots,k-1 \\ f_k''' & i=k \end{cases}$$

Then

$$E_{H_2}[\{R_{ij}\}] = \begin{Bmatrix} N_1f_1''' & N_1S_1f_2 & \cdots & N_1S_1\cdots S_{k-1}f_k \\ & N_2f_2''' & \cdots & N_2S_2\cdots S_{k-1}f_k \\ & & \ddots & \vdots \\ & & & N_kf_k''' \end{Bmatrix}$$

and

$$E_{H_2}[\{Q_{ij}\}] = E_{H_1}[\{Q_{ij}\}] \ .$$

The likelihood under H_2 is given by

$$P_{H_2}[\{R_{ij}\},\{Q_{ij}\}] = \prod_{i=1}^{k} \binom{N_i}{R_{ii},\ldots,R_{ik}}\binom{M_i}{Q_{ii},\ldots,Q_{ik}} f_i^{W_i-R_{ii}}$$

$$S_i^{Z_{i+1}} f_i'''^{R_{ii}} f_i'^{Q_{ii}} S_i'^{Q_{i\cdot}-Q_{ii}}$$

$$(1-\rho_i''')^{N_i-R_{i\cdot}} (1-\rho_i)^{M_i-Q_{i\cdot}}$$

Variances and covariances of maximum likelihood estimators under H_2

Exact variances of $\hat{f}_i''',\hat{f}_i',\widehat{S_{k-1}f_k},\widehat{S'_{k-1}f_k}$ are the appropriate binomial variances. Asymptotic variances of the other estimators are:

$$\mathrm{Var}\ (\hat{f}_i) \doteq f_i^2\left[\frac{1}{E(R_{i\cdot}-R_{ii})}-\frac{1}{N_i}-\frac{1}{E(Z_{i+1}-R_{i\cdot}+R_{ii})}+\frac{1}{E(W_i-R_{ii})}\right]$$

$$\mathrm{Var}\ (\hat{S}_i) \doteq S_i^2\left[\frac{1}{E(R_{i\cdot}-R_{ii})}-\frac{1}{N_i}+\frac{1}{E(R_{i+1\cdot}-R_{i+1,i+1})}-\frac{1}{N_{i+1}}\right.$$
$$\left.+\frac{1}{E(Z_{i+2}-R_{i+1\cdot}+R_{i+1,i+1})}-\frac{1}{E(W_{i+1}+Z_{i+2}-R_{i+1\cdot})}\right]$$

$$\mathrm{Var}\ (\hat{S}_i') \doteq S_i'^2\left[\frac{1}{E(Q_{i\cdot}-Q_{ii})}-\frac{1}{M_i}+\frac{1}{E(R_{i+1\cdot}-R_{i+1,i+1})}-\frac{1}{N_{i+1}}\right.$$
$$\left.+\frac{1}{E(Z_{i+2}-R_{i+1\cdot}+R_{i+1,i+1})}-\frac{1}{E(W_{i+1}+Z_{i+2}-R_{i+1\cdot})}\right]$$

Non-zero asymptotic covariances are:

$$\mathrm{Cov}\ (\hat{f}_i''',\hat{f}_i) \doteq -f_i'''f_i/N_i$$

$$\mathrm{Cov}\ (\hat{f}_i''',\hat{S}_i) \doteq -f_i'''S_i/N_i$$

$$\mathrm{Cov}\ (\hat{f}_{i+1}''',\hat{S}_i) \doteq f_{i+1}'''S_i/N_{i+1}$$

$$\mathrm{Cov}\ (\hat{f}_{i+1}''',\hat{S}_i') \doteq f_{i+1}'''S_i'/N_{i+1}$$

$$\mathrm{Cov}\ (\hat{f}_i',\hat{S}_i') \doteq -f_i'S_i'/M_i$$

$$\mathrm{Cov}\ (\hat{f}_i,\hat{S}_i) \doteq f_iS_i\left[\frac{1}{E(R_{i\cdot}-R_{ii})}-\frac{1}{N_i}\right]$$

$$\text{Cov}\,(\hat{f}_{i+1},\hat{S}_i) \doteq -f_{i+1}S_i\left[\frac{1}{E(R_{i+1.}-R_{i+1,i+1})} - \frac{1}{N_{i+1}}\right]$$

$$\text{Cov}\,(\hat{f}_{i+1},\hat{S}'_i) \doteq -S'_i f_{i+1}\left[\frac{1}{E(R_{i+1.}-R_{i+1,i+1})} - \frac{1}{N_{i+1}}\right]$$

$$\text{Cov}\,(\hat{S}_i,\hat{S}_{i+1}) \doteq -S_i S_{i+1}\left[\frac{1}{E(R_{i+1.}-R_{i+1,i+1})} - \frac{1}{N_{i+1}}\right]$$

$$\text{Cov}\,(\hat{S}_i,\hat{S}'_i) \doteq S_i S'_i\left[\frac{1}{E(R_{i+1.}-R_{i+1,i+1})} - \frac{1}{N_{i+1}} + \frac{1}{E(Z_{i+2}-R_{i+1.}+R_{i+1,i+1})} - \frac{1}{E(W_{i+1}+Z_{i+2}-R_{i+1.})}\right]$$

$$\text{Cov}\,(\hat{S}_{i+1},\hat{S}'_i) \doteq -S_{i+1}S'_i\left[\frac{1}{E(R_{i+1.}-R_{i+1,i+1})} - \frac{1}{N_{i+1}}\right].$$

All expectations are with respect to the model under H_2.

5. The Model Under H_3

Under the hypotheses H_1 and H_2 it is assumed that survival rates are age-specific for birds less than one year old, but independent of age for all other birds. It may be, however, that survival rates are also age-dependent for yearlings, and independent of age only for birds more than two years old. This leads to a third modification of the basic model, and to the hypothesis H_3, which assumes that survival rates are age-specific for young-of-the-year birds and yearlings only.

In most cases it is possible to distinguish between young-of-the-year birds and adults only, so that the experimental situation remains the same, with data collected for these two classes of birds only. Each batch of adults released will include yearlings and older birds in unknown proportions. So in a given year the survival rate for a newly released batch of adults will be different from that of birds released in previous years. Therefore, new parameters f_i''', S_i''' are defined by:

f_i''' = P [adult released in year i is captured and reported in year i], $i=1,\ldots,k$

S_i''' = P [adult released in year i survives year i], $i=1,\ldots,k-1$.

Also

f_i'' = P [yearling released as young-of-the-year in year i-1 is captured and reported in year i given it survived year i-1], $i=2,\ldots,k$

S_i'' = P [yearling released as young-of-the-year in year i-1 survives year i given it survived year i-1], $i=2,\ldots,k-1$.

$$E_{H_3}\left[\frac{Q_{i\cdot}}{M_i}\right] = \rho_i' = \begin{cases} f_i' + S_i'f_{i+1}'' + S_i'S_{i+1}''\rho_{i+2} = f_i'+S_i'\rho_{i+1}'', & i=1,\ldots,k-2 \\ f_{k-1}' + S_{k-1}'f_k'' = f_{k-1}'+S_{k-1}'\rho_k'' & i=k-1 \\ f_k' & i=k \end{cases}$$

$$E_{H_3}\left[\frac{R_{i\cdot}}{N_i}\right] = \rho_i''' = \begin{cases} f_i''' + S_i'''\rho_{i+1}, & i=1,\ldots,k-1 \\ f_k''' & i=k \end{cases}$$

where f_i, S_i, f_i', S_i' are as defined under H_2.

Then the hypothesis H_3 can be characterized by:

$$E_{H_3}[\{R_{ij}\}] = \begin{Bmatrix} N_1f_1''' & N_1S_1'''f_2 & N_1S_1'''S_2f_3 & \cdots & N_1S_1'''\cdots S_{k-1}f_k \\ & N_2f_2''' & N_2S_2'''f_3 & \cdots & N_2S_2'''\cdots S_{k-1}f_k \\ & & N_3f_3''' & \cdots & N_3S_3'''\cdots S_{k-1}f_k \\ & & & \ddots & \vdots \\ & & & & N_kf_k''' \end{Bmatrix}$$

$$E_{H_3}[\{Q_{ij}\}] = \begin{Bmatrix} M_1f_1' & M_1S_1'f_2'' & M_1S_1'S_2''f_3 & \cdots & M_1S_1'S_2''\cdots S_{k-1}f_k \\ & M_2f_2' & M_2S_2'f_3'' & \cdots & M_2S_2'S_3''\cdots S_{k-1}f_k \\ & & M_3f_3' & \cdots & M_3S_3'S_4''\cdots S_{k-1}f_k \\ & & & \ddots & \vdots \\ & & & & M_kf_k' \end{Bmatrix}$$

Again assuming that banded birds released into the population suffer independent fates, we obtain:

$$P_{H_3}[\{R_{ij}\},\{Q_{ij}\}] = \prod_{i=1}^{k} \binom{N_i}{R_{ii},\ldots,R_{ik}}\binom{M_i}{Q_{ii},\ldots,Q_{ik}} f_i'''^{R_{ii}} S_i'''^{R_{i.}-R_{ii}}$$

$$f_i'^{Q_{ii}} S_i'^{Q_{i.}-Q_{ii}} (1-\rho_i''')^{N_i-R_{i.}} (1-\rho_i')^{M_i-Q_{i.}}$$

$$\prod_{i=2}^{k} f_i''^{Q_{i-1,i}} S_i''^{Q_{i-1.}-Q_{i-1,i-1}-Q_{i-1,i}} f_i^{W_i-R_{ii}-Q_{i-1,i}}$$

$$\prod_{i=2}^{k} S_i^{Z_{i+1}-R_{i.}+R_{ii}-Q_{i-1.}+Q_{i-1,i-1}+Q_{i-1,i}}$$

The minimal sufficient statistic is $\mathscr{S}_3 = \{R_{1.},\ldots,R_{k.}, Q_{1.},\ldots,Q_{k.}, R_{11},\ldots,R_{k-1,k-1}, Q_{11},\ldots,Q_{k-1,k-1}, Q_{12},Q_{23},\ldots,Q_{k-2,k-1}, T_3+U_3,\ldots,T_k+U_k\} = \{\mathscr{S}_2,Q_{12},Q_{23},\ldots,Q_{k-2,k-1}\}$.

Under H_3, parameters of interest are not identifiable and hence not estimable. However the assumptions of the hypothesis H_3 are of interest and so tests related to H_3 are developed in the next sections. In other words, tests between H_1 and H_3 and between H_2 and H_3 are of interest because rejection of H_1 and H_2 in favor of H_3 would imply that survival and reported exploitation rates are age-specific for at least the first two years of life.

The model under H_3 is the last in our sequence of models for the experimental situation where data are recorded separately for young-of-the-year and adult birds.

6. Tests to Discriminate Between H_1, H_2, H_3

The hypotheses H_1, H_2, H_3 are such that $H_1 \subset H_2 \subset H_3$ and there is also a relationship between the corresponding minimal sufficient statistics $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_3$ given by $\mathscr{S}_2 = \{\mathscr{S}_1, R_{22},\ldots,R_{k-1,k-1}\}$, $\mathscr{S}_3 = \{\mathscr{S}_2, Q_{12},Q_{23},\ldots,Q_{k-2,k-1}\}$. This suggests the construction of conditional tests $\phi(\mathscr{S}_1,\mathscr{S}_2)$ and $\phi(\mathscr{S}_2,\mathscr{S}_3)$ of H_1 vs H_2, and H_2 vs H_3 respectively, such that $E_{H_1}[\phi(\mathscr{S}_1,\mathscr{S}_2)|\mathscr{S}_1] = \alpha$, $E_{H_2}[\phi(\mathscr{S}_2,\mathscr{S}_3)|\mathscr{S}_2] = \alpha$. Actually $\phi(\mathscr{S}_2,\mathscr{S}_3)$ will be a test of $H_1 \cup H_2$ vs H_3 since $E_{H_1}[\phi(\mathscr{S}_2,\mathscr{S}_3)|\mathscr{S}_2] = E_{H_2}[\phi(\mathscr{S}_2,\mathscr{S}_3)|\mathscr{S}_2]$.

The tests $\phi(\mathscr{S}_1,\mathscr{S}_2)$, $\phi(\mathscr{S}_2,\mathscr{S}_3)$ are based on the conditional distributions given by $P_{H_1}[\mathscr{S}_2|\mathscr{S}_1]$, $P_{H_2}[\mathscr{S}_3|\mathscr{S}_2]$ respectively. The derivations of these distributions are analogous to the derivation of $P_{H_0}[\mathscr{S}_1|\mathscr{S}_0]$ outlined in section 3, and we obtain

$$P_{H_1}[\mathscr{S}_2|\mathscr{S}_1] = \prod_{i=2}^{k-1} \frac{\binom{R_{i\cdot}}{R_{ii}}\binom{T_i+U_i-Q_{i\cdot}-R_{i\cdot}}{W_i-R_{ii}}}{\binom{T_i+U_i-Q_{i\cdot}}{W_i}} ,$$

$$P_{H_2}[\mathscr{S}_3|\mathscr{S}_2] = \prod_{i=2}^{k-1} \frac{\binom{Q_{i-1\cdot}-Q_{i-1,i-1}}{Q_{i-1,i}}\binom{T_i+U_i-R_{i\cdot}-Q_{i\cdot}-Q_{i-1\cdot}+Q_{i-1,i-1}}{W_i-R_{ii}-Q_{i-1,i}}}{\binom{T_i+U_i-R_{i\cdot}-Q_{i\cdot}}{W_i-R_{ii}}}$$

Both of the above expressions are in the form of a product of k-2 terms, each term being a hypergeometric probability, and the corresponding test statistics are obtained using the usual chi-square approximation as outlined in section 3. Thus the test of H_1 vs H_2 is based on a test statistic which is asymptotically chi-square on k-2 degrees of freedom under H_1, and similarly the test statistic for H_2 vs H_3 is asymptotically chi-square on k-2 degrees of freedom under H_2.

7. Non-Discriminant Goodness of Fit Tests of the Models Under H_1, H_2, H_3

Robson and Youngs (1971) developed a size α goodness of fit test of their model based on the conditional distribution of the data array given the minimal sufficient statistic. Analogous goodness of fit tests can be developed for the models under H_1, H_2, H_3, based on the residual distributions given by $P_{H_1}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_1]$, $P_{H_2}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_2]$, $P_{H_3}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{S}_3]$ respectively.

The derivation of each of these distributions is analogous to that given by Robson and Youngs (1971) and is omitted here.

Define

$$R^*_{ij} = \begin{cases} R_{1j} & i=1 \\ R_{1j}+R_{2j}+\ldots+R_{ij} & i=2,\ldots,j, \end{cases}$$

$$Q^*_{ij} = \begin{cases} Q_{1j} & i=1 \\ Q_{1j}+Q_{2j}+\ldots+Q_{ij} & i=2,\ldots,j\ . \end{cases}$$

Then $P_{H_1}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{A}_1] =$

$$\prod_{i=2}^{k-1} \frac{\binom{Z_i}{R^*_{i-1,i}+Q^*_{i-2,i},\ldots,R^*_{i-1,k}+Q^*_{i-2,k}}\binom{R_{i\cdot}}{R_{ii},\ldots,R_{ik}}\binom{Q_{i-1\cdot}-Q_{i-1,i-1}}{Q_{i-1,i},\ldots,Q_{i-1,k}}}{\binom{T_i+U_i-Q_{i\cdot}}{R^*_{ii}+Q^*_{i-1,i},\ldots,R^*_{ik}+Q^*_{i-1,k}}},$$

$$P_{H_2}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{A}_2] = \frac{\binom{R_{1\cdot}-R_{11}}{R_{12},\ldots,R_{1k}}\binom{Q_{1\cdot}-Q_{11}}{Q_{12},\ldots,Q_{1k}}}{\binom{T_2+U_2-R_{2\cdot}-Q_{2\cdot}}{R_{12}+Q_{12},\ldots,R_{1k}+Q_{1k}}}$$

$$\prod_{i=2}^{k-2} \frac{\binom{T_{i+1}+U_{i+1}-R_{i+1\cdot}-Q_{i+1\cdot}-R_{i\cdot}+R_{ii}-Q_{i\cdot}+Q_{ii}}{R^*_{i-1,i+1}+Q^*_{i-1,i+1},\ldots,R^*_{i-1,k}+Q^*_{i-1,k}}\binom{R_{i\cdot}-R_{ii}}{R_{i,i+1},\ldots,R_{ik}}\binom{Q_{i\cdot}-Q_{ii}}{Q_{i,i+1},\ldots,Q_{ik}}}{\binom{T_{i+1}+U_{i+1}-R_{i+1\cdot}-Q_{i+1\cdot}}{R^*_{i,i+1}+Q^*_{i,i+1},\ldots,R^*_{i,k}+Q^*_{i,k}}},$$

and $P_{H_3}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{A}_3] =$

$$\prod_{i=2}^{k-2} \frac{\binom{Z_{i+1}-R_{i\cdot}+R_{ii}-Q_{i-1\cdot}+Q_{i-1,i-1}+Q_{i-1,i}}{R^*_{i-1,i+1}+Q^*_{i-2,i+1},\ldots,R^*_{i-1,k}+Q^*_{i-2,k}}\binom{R_{i\cdot}-R_{ii}}{R_{i,i+1},\ldots,R_{ik}}\binom{Q_{i-1\cdot}-Q_{i-1,i-1}-Q_{i-1,i}}{Q_{i-1,i+1},\ldots,Q_{i-1,k}}}{\binom{Z_{i+1}}{R^*_{i,i+1}+Q^*_{i-1,i+1},\ldots,R^*_{ik}+Q^*_{i-1,k}}}$$

where $Q_{0,i} \equiv 0$.

For $m = 1,2,3$, $P_{H_m}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{A}_m]$ is a product of multihypergeometric distributions, and a chi-square test can be applied to the contingency table corresponding to each multihypergeometric term. For example, for $m = 1$ the appropriate contingency tables are

$$\boxed{\begin{array}{llll} R_{ii} & R_{i,i+1} & \cdots & R_{ik} \\ Q_{i-1,i} & Q_{i-1,i+1} & \cdots & Q_{ik} \\ R^*_{i-1,i}+Q^*_{i-2,i} & R^*_{i-1,i+1}+Q^*_{i-2,i+1} & \cdots & R^*_{i-1,k}+Q^*_{i-2,k} \end{array}} ,$$

each yielding a chi-square statistic on 2(k-i) degrees of freedom for $i=2,\ldots,k-1$. The chi-square statistics obtained in this way for a given distribution $P_{H_m}[\{R_{ij}\},\{Q_{ij}\}|\mathcal{S}_m]$ are asymptotically independent and may be added to give a single statistic on which the goodness of fit test to H_m is based. For example, the goodness of fit test to H_1 is based on a chi-square statistic with degrees of freedom equal to $\sum_{i=2}^{k-1} 2(k-i) = (k-2)(k-1)$.

8. Discussion of Tests Derived in Sections 3, 6, 7

8.1 Comments related to the statistics $\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3$

Considerable use is made of the sufficient statistics $\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3$ in deriving the tests above, and so it should be noted that there other models for which the statistics $\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3$ are minimal sufficient. For example, $\mathcal{S}_1$ is minimal sufficient for models which are identical to the model under H_1 except that $E[Q_{i.}/M_i] = \rho'_i$ is defined by $\rho'_i = f_i+S'_i\rho_{i+1}$ or by $\rho'_i = f'_i+S_i\rho_{i+1}$, for each i in any subset of the index set $\{1,2,\ldots,k-1\}$. The dimension of the natural parameter space corresponding to the model under H_1 is unaffected if $S'_i = S_i$ or if $f'_i = f_i$, $i\neq k$, but the number of biologically meaningful parameters is reduced.

In the context of this study, the model under H_1 has a meaningful biological interpretation but the other models described above do not. This is because a difference in the harvesting rate of young birds should be reflected by a difference in their survival rate, and vice versa. Also young and adult birds released in the same year should have the same reporting rate.

The situation is more complex for the statistics $\mathscr{J}_2, \mathscr{J}_3$, since analogous changes in the definition of ρ_i''', ρ_i'' also give rise to new models. Again it is difficult to find a meaningful biological interpretation for any of these models except the models under hypotheses H_2, H_3, and so all except these last two models have been largely ignored.

It should now be recognized that tests based on $\mathscr{J}_1, \mathscr{J}_2, \mathscr{J}_3$ are in fact related to the composite hypotheses comprising the numerous alternative models described above. However, in view of the biological considerations outlined above, tests based on $\mathscr{J}_1, \mathscr{J}_2, \mathscr{J}_3$ will be regarded as being related to just those models under hypotheses H_1, H_2 and H_3.

8.2 Some asymptotic considerations

Define $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{J}_m)$ to be the test statistic associated with $P_{H_m}[\{R_{ij}\},\{Q_{ij}\}|\mathscr{J}_m]$, m=1,2,3, as described in section 7.

Similarly, define $\chi^2(\mathscr{J}_{m+j}; \mathscr{J}_m)$ to be the test statistic associated with $P_{H_m}(\mathscr{J}_{m+j}|\mathscr{J}_m)$, m=1,2, j=1,...,(3-m).

We note the following

(i) Asymptotically, $\chi^2(X; \mathscr{J}_m)$, (where $X = (\{R_{ij}\},\{Q_{ij}\})$ or $\mathscr{J}_{m+j}$), is a sum of independent central chi-square variables under H_ℓ for $\ell \leq m$, and is a sum of independent non-central chi-square variables under H_ℓ for $\ell > m$.

Approximate power for the statistics $\chi^2(X; \mathscr{J}_j)$ can be obtained using methods outlined by Meng and Chapman (1966). They obtained asymptotic power results, using alternatives of Pitman type, for chi-square tests for contingency tables. Evaluation of their formulae for non-centrality parameters in the "comparative trials" situation involves making the correct substitution for their $Q_i = N_i/N$, where N_i are the contingency table row totals, and N the grand total. Row totals and the grand total for the contingency tables associated with the statistics $\chi^2(X; \mathcal{S}_j)$ are components of the sufficient statistic $\mathcal{S}_j$ and so Q_i is replaced by the stochastic limit of the appropriate ratios of the components of $\mathcal{S}_j$. Motivation for this can be found in results of Harkness and Katz (1964).

(ii) The order relationships between $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$ imply

(a) $P_{H_j}[\{R_{ij}\},\{Q_{ij}\}|\mathcal{S}_1] = P_{H_j}[\mathcal{S}_2|\mathcal{S}_1]\ P_{H_j}[\{R_{ij}\},\{Q_{ij}\}|\mathcal{S}_2]$,

(b) $P_{H_j}[\mathcal{S}_3|\mathcal{S}_1] = P_{H_j}[\mathcal{S}_2|\mathcal{S}_1]\ P_{H_j}[\mathcal{S}_3|\mathcal{S}_2]$, $\qquad j=1,2,3$

(iii) The degrees of freedom for the two statistics associated with the right-hand side of each of (a), (b) add to the degrees of freedom of the statistic associated with the corresponding left-hand side. For example, the statistics $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathcal{S}_1)$, $\chi^2(\mathcal{S}_2; \mathcal{S}_1)$, $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathcal{S}_2)$ have degrees of freedom $(k-2)(k-1)$, $(k-2)$, $(k-2)^2$ respectively and $k-2 + (k-2)^2 = (k-2)(k-1)$.

(iv) For $j=2,3$, if the non-centrality parameters for the statistics associated with each of (a), (b) are evaluated for Pitman type alternatives contained in H_2 or H_3, using the Meng and Chapman formulae as described above, then these non-centrality parameters satisfy the same additive relationship described in (iii) for the corresponding degrees of freedom.

(v) If X_i is a non-central chi-square variable with degrees of freedom v_i and non-centrality parameter λ_i, $i=1,2$, and X_1, X_2 are independent, then X_1+X_2 is a non-central chi-square variable with degrees of freedom v_1+v_2 and non-centrality parameter $\lambda_1+\lambda_2$.

These observations suggest (but do not prove) asymptotic equivalence of the statistics $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_1)$ and $\chi^2(\mathscr{I}_2; \mathscr{I}_1) + \chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_2)$ and also of the statistics $\chi^2(\mathscr{I}_3; \mathscr{I}_1)$ and $\chi^2(\mathscr{I}_2; \mathscr{I}_1) + \chi^2(\mathscr{I}_3; \mathscr{I}_2)$. This result would be useful both for practical reasons involved in the computation of the goodness of fit tests, and for the purpose of making power comparisons.

The statistics $\chi^2(\mathscr{I}_{m+1}; \mathscr{I}_m)$ are relatively easy to compute in comparison to the statistics $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_m)$. Thus, for large enough N_i, M_i, and if the above conjecture is true, the goodness of fit test to H_2 could be computed as the sum $\chi^2(\mathscr{I}_3; \mathscr{I}_2) + \chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_3)$ and the goodness of fit test to H_1 as $\chi^2(\mathscr{I}_2; \mathscr{I}_1) + \chi^2(\mathscr{I}_3; \mathscr{I}_2) + \chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_3)$, making it necessary to compute $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_m)$ for m = 3 only.

Asymptotic equivalence of $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_1)$ and $\chi^2(\mathscr{I}_2; \mathscr{I}_1) + \chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_2)$, which depends on the asymptotic independence of $\chi^2(\mathscr{I}_2; \mathscr{I}_1)$, $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_2)$, would imply that, asymptotically under H_2, the statistics $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_1)$, $\chi^2(\mathscr{I}_2; \mathscr{I}_1)$ have the same non-centrality parameter. The difference in degrees of freedom of these two statistics would then suggest that for making tests about H_1, the statistic $\chi^2(\mathscr{I}_2; \mathscr{I}_1)$ would be more powerful against alternatives contained in H_2, than the statistic $\chi^2(\{R_{ij}\},\{Q_{ij}\}; \mathscr{I}_1)$, or the (probably) asymptotically equivalent conventional chi-square goodness of fit statistic.

9. Applications and Extensions

So far we have discussed the application of the methods of this paper to band return data from exploited bird populations only. However there are data from other types of populations to which these methods can be applied. For example, blackbird populations, though not exploited, are subjected to yearly banding operations and substantial numbers of banded birds are found dead each year and reported to the banding office. Also models under H_1, H_3 are applicable to other exploited animal populations, whenever there is a yearly banding program in which young animals are distinguished from adults. Many fish populations are of this type.

The methods of this paper are easily extended to experimental situations where data are recorded separately for more than two age groups. This is the case in some tagging operations carried out in fish populations. If data for more than two age groups are available, then models can be developed which permit even greater flexibility in the assumptions about annual survival, exploitation, and reporting rates. A sequence of models, estimation formulae and tests have been derived for the situation where three age groups are recognizable, namely young-of-the-year, yearlings, and animals more than two years old. Details are given in Brownie (1973). These three age groups can be distinguished in some species of geese.

In the derivations above, it has been assumed for simplicity that the data arrays $\{R_{ij}\},\{Q_{ij}\}$ are triangular, i.e., that returns are recorded only for as many years as banded birds are released. In practice it is common for returns to be recorded for several years after the last batch is released, and in this case the arrays $\{R_{ij}\},\{Q_{ij}\}$ have more columns than rows. This situation is referred to briefly at the end of sections 2, 3 and details can be found in Brownie (1973). A numerical example for data of this type follows.

10. Numerical Example

The above methods are illustrated using data for mallards banded in southwestern Saskatchewan during 1956 to 1959. Estimates and tests are presented in Tables 1 to 7 in the form of output from a fortran program developed by Mrs. E. Keokosky of the Computer Activities Group at Cornell University.

In Table 1 the "input matrices" for adults and young correspond to the $\{R_{ij}\}$ and $\{Q_{ij}\}$ arrays respectively. The first column gives the year of banding and the second column gives the number of birds banded. The "Basic parameters" correspond to the summary statistics $R_{i.}$, $R_{.i}$, $Q_{i.}$, $Q_{.i}$ etc., (reading from left to right).

Tables 2 and 4 contain estimates and confidence intervals for the models under hypotheses H_1 and H_2 respectively. In these tables FHAT should be read as $\hat{f}$, F'HAT as $\hat{f}'$ etc.

Tables 3 and 5 contain the estimated covariances under H_1 and H_2 respectively. Here FHATP(I) should be read as $\hat{f}_i'$, FHAT3P(I) as $\hat{f}_i'''$ etc.

Table 6 contains the tests of H_0 vs H_1, H_1 vs H_2 and H_2 vs H_3, and Table 7 contains the goodness of fit tests. In the goodness of fit tests some pooling across columns in the contingency tables may be necessary in order to justify applying a chi-square test. The fortran program uses an algorithm for pooling columns which is based on a procedure described by Robson (1971). Thus in the first contingency table in the goodness of fit test to H_3 pooling has resulted in a table with 5 columns instead of 7 and the degrees of freedom are therefore 8 instead of 12. The goodness of fit tests to H_1, H_2 are computed using the approximate method described in section 8.2.

Table 6 shows that the test of H_0 vs H_1 is very significant and we conclude that young-of-the-year and adult birds have different survival and reported exploitation rates. None of the other tests are significant and we conclude that the model under H_1 is appropriate for this data.

Availability of the FORTRAN Program

A source deck and brief documentation for the FORTRAN program, written for an IBM 360, are available from The Migratory Bird and Habitat Research Laboratory, Patuxent Wildlife Research Center, Laurel, Maryland 20810.

Acknowledgements

This research was supported by contract No. 14-16-0008-664 from the USDI, Bureau of Sport Fisheries and Wildlife, and the authors thank Dr. D. R. Anderson of that bureau for his many helpful comments.

Dr. J. B. Gollop deserves credit for banding most of the mallards in the banding experiment used in the example of section 10.

References

Brownie, C. (1973). Stochastic models allowing age-dependent survival rates for banding experiments on exploited bird populations. Ph.D. Thesis, Biometrics Unit, Cornell University.

Harkness, W. L. and Katz, L. (1964). Comparison of the power functions for the test of independence in 2 x 2 contingency tables. Ann. Math. Stat. 35, 1115-1122.

Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration--stochastic model. Biometrika 52, 225-248.

Meng, R. C. and Chapman, D. G. (1966). The power of chi-square tests for contingency tables. J. Amer. Stat. Assoc. 61, 965-975.

Robson, D. S. (1971). A non-prejudicial method of defining class intervals for comparing frequency distribution by chi-square. Paper No. BU-366-M in the Biometrics Unit Mimeograph Series, Cornell University.

Robson, D. S. and Youngs, W. D. (1971). Statistical analysis of reported tag-recaptures in the harvest from an exploited population. Paper No. BU-369-M in the Biometrics Unit Mimeograph Series, Cornell University.

Seber, G. A. F. (1970). Estimating time-specific survival and reporting rates for adult birds from band returns. Biometrika 57, 313-318.

Key Words: Age-specific survival rates; Age-specific exploitation rates; Banding experiment; Band-return; Tag-recapture.

Table 1

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

ADULTS INPUT MATRIX

1956	578.	36.	24.	12.	5.	5.	2.	2.	2.	1.
1957	724.	0.	51.	40.	23.	17.	6.	3.	1.	6.
1958	2639.	0.	0.	185.	80.	62.	29.	19.	13.	17.
1959	2045.	0.	0.	0.	125.	89.	38.	22.	20.	21.

YOUNG INPUT MATRIX

1956	1886.	222.	81.	44.	18.	16.	5.	4.	4.	4.
1957	4597.	0.	401.	167.	63.	57.	26.	3.	12.	12.
1958	4632.	0.	0.	441.	120.	72.	35.	13.	21.	11.
1959	1351.	0.	0.	0.	95.	22.	19.	2.	5.	4.

BASIC PARAMETERS

I	RROW(I)	RCOL(I)	QROW(I)	QCOL(I)	T(I)	U(I)	W(I)	Z(I)
1	89.00	36.00	398.00	222.00	89.00	398.00	36.00	0.0
2	147.00	75.00	741.00	482.00	200.00	917.00	156.00	53.00
3	405.00	237.00	713.00	652.00	530.00	1148.00	448.00	220.00
4	315.00	233.00	147.00	296.00	608.00	643.00	434.00	517.00
					375.00	347.00		670.00

Table 2

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

ESTIMATES UNDER H1

FHAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.0623	0.0101	0.0426 - 0.0820
2	0.0842	0.0081	0.0684 - 0.1001
3	0.0712	0.0041	0.0632 - 0.0793
4	0.0606	0.0039	0.0530 - 0.0681

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.0696	0.0035	0.0627 - 0.0765

SHAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.4516	0.0678	0.3187 - 0.5846
2	0.7741	0.0750	0.6270 - 0.9212
3	0.5338	0.0402	0.4550 - 0.6126

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.5865	0.0265	0.5345 - 0.6385

F'HAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.1177	0.0074	0.1032 - 0.1323
2	0.0872	0.0042	0.0791 - 0.0954
3	0.0952	0.0043	0.0868 - 0.1037
4	0.0703	0.0070	0.0567 - 0.0840

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.0926	0.0030	0.0868 - 0.0984

S'HAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.4596	0.0473	0.3670 - 0.5522
2	0.4819	0.0334	0.4164 - 0.5475
3	0.3812	0.0299	0.3226 - 0.4398

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.4409	0.0217	0.3984 - 0.4835

SK...SK+I-1FK+I

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.0440	0.0031	0.0380 - 0.0500
2	0.0207	0.0019	0.0171 - 0.0244
3	0.0088	0.0011	0.0066 - 0.0110
4	0.0101	0.0012	0.0077 - 0.0125
5	0.0098	0.0012	0.0075 - 0.0122

SK/S'K

ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
2.4287	0.3583	1.7265 - 3.1309

Table 3

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

ESTIMATED NON-ZERO ASYSMPTOTIC COVARIANCES UNDER H1

I	YR	COVAR(FHAT(I),SHAT(I))	CORR(FHAT(I),SHAT(I))	COVAR(SHAT(I),SHAT(I+1))	CORR(SHAT(I),SHAT(I+1))
1	1956	-0.000048665	-0.071371609	-0.001895337	-0.372347718
2	1957	0.000180105	0.297486865	-0.000863670	-0.286249614
3	1958	0.000040081	0.244083910		

I	YR	COVAR(FHAT(I+1),SHAT(I))	CORR(FHAT(I+1),SHAT(I))	COVAR(FHATP(I),SHATP(I))	CORR(FHATP(I),SHATP(I))
1	1956	-0.000206255	-0.376895111	-0.000028685	-0.081795182
2	1957	-0.000115280	-0.376133510	-0.000009145	-0.065716037
3	1958	-0.000086804	-0.557912343	-0.000007836	-0.060795795

I	YR	COVAR(SHATP(I),SHAT(I))	CORR(SHATP(I),SHAT(I))	COVAR(SHATP(I),SHAT(I+1))	CORR(SHATP(I),SHAT(I+1))
1	1956	0.001125332	0.351039151	-0.001928900	-0.543889314
2	1957	0.000779786	0.310766900	-0.000537698	-0.399951034
3	1958	0.000546496	0.454782392		

I	YR	COVAR(SHATP(I),FHAT(I+1))	CORR(SHATP(I),FHAT(I+1))
1	1956	-0.000209907	-0.550531806
2	1957	-0.000071770	-0.525537757
3	1958	-0.000061996	-0.536042148

Table 4

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

ESTIMATES UNDER H2

FHAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1			
2	0.1123	0.0183	0.0764 - 0.1482
3	0.0738	0.0079	0.0584 - 0.0892
4	0.0598	0.0060	0.0480 - 0.0716

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.0820	0.0069	0.0684 - 0.0956

SHAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.3745	0.0647	0.2476 - 0.5013
2	0.8436	0.1026	0.6425 - 1.0446
3	0.5459	0.0539	0.4402 - 0.6515
4			

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.5880	0.0299	0.5293 - 0.6467

F'HAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.1177	0.0074	0.1032 - 0.1323
2	0.0872	0.0042	0.0791 - 0.0954
3	0.0952	0.0043	0.0868 - 0.1037
4	0.0703	0.0070	0.0567 - 0.0840

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.0926	0.0030	0.0868 - 0.0984

S'HAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.3811	0.0510	0.2812 - 0.4810
2	0.4705	0.0433	0.3856 - 0.5554
3	0.3845	0.0366	0.3128 - 0.4562
4			

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.4120	0.0254	0.3622 - 0.4618

F'''HAT

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.0623	0.0101	0.0426 - 0.0820
2	0.0704	0.0095	0.0518 - 0.0891
3	0.0701	0.0050	0.0604 - 0.0798
4	0.0611	0.0053	0.0507 - 0.0715
5			

AVERAGE ESTIMATE	STANDARD ERROR	95% CONFIDENCE INTERVAL
0.0660	0.0039	0.0583 - 0.0736

SK...SK+I-1FK+I

I	ESTIMATE	STANDARD DEVIATION	95% CONFIDENCE INTERVAL
1	0.0438	0.0035	0.0369 - 0.0506
2	0.0206	0.0020	0.0166 - 0.0246
3	0.0088	0.0012	0.0064 - 0.0111
4	0.0100	0.0013	0.0075 - 0.0125
5	0.0098	0.0013	0.0073 - 0.0122

Table 5

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

ESTIMATED NON-ZERO ASYSMPTOTIC COVARIANCES UNDER H2

I	YR	COVAR(FHAT3P(I),FHAT(I))	CORR(FHAT3P(I),FHAT(I))	COVAR(FHAT3P(I),SHAT(I))	CORR(FHAT3P(I),SHAT(I))
1	1956			-0.000040350	-0.062024460
2	1957	-0.000010924	-0.062700846	-0.000082075	-0.084144002
3	1958	-0.000001961	-0.050199862	-0.000014500	-0.054110327
4	1959	-0.000001788	-0.056160593		

I	YR	COVAR(FHAT3P(I+1),SHAT(I))	CORR(FHAT3P(I+1),SHAT(I))	COVAR(FHAT3P(I+1),SHATP(I))	CORR(FHAT3P(I+1),SHATP(I))
1	1956	0.000036433	0.059194816	0.000037078	0.076500967
2	1957	0.000022408	0.043958577	0.000012499	0.058069756
3	1958	0.000016316	0.057122705	0.000011493	0.059308664

I	YR	COVAR(FHAT(I),SHAT(I))	CORR(FHAT(I),SHAT(I))	COVAR(SHAT(I),FHAT(I+1))	CORR(SHAT(I),FHAT(I+1))
1	1956			-0.000379884	-0.320398482
2	1957	0.000855790	0.455438486	-0.000259462	-0.321856414
3	1958	0.000167897	0.396185615	-0.000155869	-0.481085192

I	YR	COVAR(FHAT(I+1),SHATP(I))	CORR(FHAT(I+1),SHATP(I))	COVAR(SHAT(I),SHAT(I+1))	CORR(SHAT(I),SHAT(I+1))
1	1956	-0.000386611	-0.414069781	-0.002854097	-0.429971890
2	1957	-0.000144725	-0.425175840	-0.001918574	-0.346928398
3	1958	-0.000109794	-0.499495163		

I	YR	COVAR(SHAT(I),SHATP(I))	CORR(SHAT(I),SHATP(I))	COVAR(SHAT(I+1),SHATP(I))	CORR(SHAT(I+1),SHATP(I))
1	1956	0.001817028	0.550896537	-0.002904634	-0.555678153
2	1957	0.002281441	0.513616070	-0.001070163	-0.458295784
3	1958	0.001173296	0.594878172		

I	YR	COVAR(FHATP(I),SHATP(I))	CORR(FHATP(I),SHATP(I))
1	1956	-0.000023784	-0.062890847
2	1957	-0.000008929	-0.049538005
3	1958	-0.000007903	-0.050099896

Table 6

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

CONTINGENCY CHI-SQUARE TEST OF H0 VS H1

	2 X 2 CONTINGENCY TABLE		CORRESPONDING CHI-SQUARE STATISTIC WITH 1 DEGREE OF FREEDOM	2 X 2 CONTINGENCY TABLE		CORRESPONDING CHI-SQUARE STATISTIC WITH 1 DEGREE OF FREEDOM
	R(I.)	N(I)-R(I.)		W(I)	Z(I+1)	
	Q(I.)	M(I)-Q(I.)		Q(I,I)	Q(I.)-Q(I,I)	
I= 1	89	489		36	53	
	398	1488	9.080	222	176	6.861
I= 2	147	577		156	220	
	741	3856	7.878	401	340	15.907
I= 3	405	2234		448	517	
	713	3919	0.003	441	272	39.171
I= 4	315	1730		434	670	
	147	1204	14.158	95	52	34.063

TOTAL CHI-SQUARE WITH 8 DEGREES OF FREEDOM = 127.120

CONTINGENCY CHI-SQUARE TEST OF H1 VS H2

	2 X 2 CONTINGENCY TABLE		CORRESPONDING CHI-SQUARE STATISTIC WITH 1 DEGREE OF FREEDOM
I= 2	51	96	
	105	124	4.591
I= 3	185	220	
	263	297	0.156
I= 4	125	190	
	309	480	0.025

TOTAL CHI-SQUARE WITH 3 DEGREES OF FREEDOM = 4.773

CONTINGENCY CHI-SQUARE TEST OF H2 VS H3

	2 X 2 CONTINGENCY TABLE		CORRESPONDING CHI-SQUARE STATISTIC WITH 1 DEGREE OF FREEDOM
I= 1	81	95	
	24	29	0.009
I= 2	167	173	
	96	124	1.611
I= 3	120	152	
	189	328	4.276
I= 4	22	30	
	318	352	0.515

TOTAL CHI-SQUARE WITH 4 DEGREES OF FREEDOM = 6.411

Table 7

MALLARDS BANDED IN SOUTHWESTERN SASKATCHEWAN, 1956 TO 1959.

CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H3

	CONTINGENCY TABLES	CORRESPONDING CHI-SQUARE STATISTICS AND DEGREES OF FREEDOM
I= 2		
	4. 12. 17. 23. 40.	
	8. 9. 16. 18. 44.	
	4. 3. 5. 5. 12.	4.64 WITH 8 D.F.
I= 3		
	19. 13. 17. 29. 62. 80.	
	3. 12. 12. 26. 57. 63.	
	9. 7. 11. 13. 38. 46.	10.60 WITH 10 D.F.
I= 4		
	22. 21. 20. 38. 89.	
	13. 11. 21. 35. 72.	
	31. 40. 32. 68.157.	5.32 WITH 8 D.F.
I= 5		
	2. 4. 5. 19.	
	66. 72. 73.141.	6.77 WITH 3 D.F.

TOTAL CHI-SQUARE 27.32 WITH 29 D.F.

CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H1

TOTAL CHI-SQUARE 38.51 WITH 36 D.F.

CHI-SQUARE GOODNESS OF FIT TEST OF THE MODEL UNDER H2

TOTAL CHI-SQUARE 33.73 WITH 33 D.F.